Technical Note 1967-6

M. Athans

On the Equivalence of Linearized Kalman Filters

11 January 1967

Prepared for the Advanced Research Projects Agency
under Electronic Systems Division Contract AF 19(628)-5167 by

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Lexington, Massachusetts

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. This research is a part of Project DEFENDER, which is sponsored by the U.S. Advanced Research Projects Agency of the Department of Defense; it is supported by ARPA under Air Force Contract AF 19(628)-5167 (ARPA Order 498).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

ON THE EQUIVALENCE OF LINEARIZED KALMAN FILTERS

M. ATHANS

Consultant, Group 42

TECHNICAL NOTE 1967-6

11 JANUARY 1967

Research completed 24 September 1966

LEXINGTON MASSACHUSETTS

ABSTRACT

The purpose of this note is to examine the relationships of the Kalman filters which can be used to estimate the linearized state of two different state-variable representations of the same nonlinear system. It is shown that, to within first order effects, the choice of the coordinate system for the construction of the Kalman filter is immaterial.

Accepted for the Air Force
Franklin C. Hudson
Chief, Lincoln Laboratory Office

1. INTRODUCTION

It is often desirable to estimate the state of a re-entry vehicle, satellite, etc. by observing certain output signals, using a radar, in the presence of noise. In general, the equations of motion of the system are nonlinear; the nonlinear terms arise due to the inverse-square law and the nonlinear dependence of the drag force on the vehicle velocity.

In such problems, the fact that one is using a radar to make observations fixes the coordinate system in which the output signals are measured. Thus, one usually observes

a) range,
b) azimuth, and
c) elevation

in the presence of measurement noise, which is in general assumed to be white.

The prediction of the future system trajectory necessitates the estimation of the remaining state variables of the system. The estimation of these variables necessitates the definition of a coordinate system. For example, for a radar-based coordinate system we must estimate

a) range	d) range-rate
b) azimuth	e) azimuth-rate
c) elevation	f) elevation-rate

On the other hand, it may be desirable to estimate the state variables using a different coordinate system. For example, one may wish to use an inertial coordinate system or a vehicle-based coordinate system. The choice of the coordinate system fixes the (nonlinear) differential equations of the chosen state variables (equations of motion) and the transformations linking the state variables to the output variables (range, azimuth, and elevation).

Since we are dealing with one and the same physical system, the various equations are related. In point of fact, given any one state variable representation one can

a) obtain, via dynamic linearization techniques, a set of

linear perturbation differential equations, and

b) construct the corresponding Kalman filter for the estimation of the perturbation variables.

It follows that each coordinate system yields a different linearized Kalman filter. The natural question that arises is: is there any advantage in choosing any one particular state variable representation?

It is shown in this report that if the linearized equations adequately represent the system behavior, then any choice of a coordinate system yields the same answer. To accomplish this we proceed as follows. In section 2 we define two different state variable representations of the same system and with the same output variables. In section 3 we state the mathematical relations which link the two state variable representations. In section 4 we obtain the linearized perturbation equations by linearizing about a nominal trajectory. In section 5 we derive the equations of the Kalman filters which can be used to estimate the perturbations of the state from the nominal trajectory; we then show that it is immaterial, to within first-order terms, which state variable representation is used to construct the Kalman filter.

2. STATEMENT OF THE PROBLEM

Let us suppose that we are given a nonlinear, time-invariant, dynamical system S_1 with state vector $\underline{x}$ and output vector $\underline{y}$ related by the equations

$$S_1 \begin{cases} \dot{\underline{x}} = \underline{f}_1(\underline{x}) & (2.1) \\ \underline{y} = \underline{h}_1(\underline{x}) & (2.2) \end{cases}$$

We suppose that this system is completely observable.

Now let us suppose that we represent the same system by using a different coordinate system in which the state vector is $\underline{z}$ and the output vector is still $\underline{y}$. We suppose that the vectors $\underline{z}$ and $\underline{x}$ are related by

$$\underline{z} = \underline{g}(\underline{x}) \tag{2.3}$$

such that $\underline{g}(.)$ is differentiable, one-to-one, and onto for all $\underline{x}$. Thus, we suppose that for every $\underline{x}$ there is a unique $\underline{z}$ and vice-versa. In other words, the inverse transformation $\underline{g}^{-1}(.)$ exists and

$$\underline{x} = \underline{g}^{-1}(\underline{z}) \tag{2.4}$$

is well defined, for all $\underline{x}$ and $\underline{z}$.

In the new coordinate system the dynamical system is described by the relations

$$S_2 \begin{cases} \dot{\underline{z}} = \underline{f}_2(\underline{z}) & (2.5) \\ \underline{y} = \underline{h}_2(\underline{z}) & (2.6) \end{cases}$$

We next suppose that we cannot observe the output vector $\underline{y}(t)$ in the absence of noise. So we let $\underline{v}(t)$ denote a vector-valued white noise process, with zero mean and covariance matrix $\underline{R}(t)$, i. e.

$$E\left\{\underline{v}(t)\right\} = \underline{0} \tag{2.7}$$

$$\mathrm{cov}\left[\underline{v}(t);\ \underline{v}(\tau)\right] = E\left\{\underline{v}(t)\underline{v}'(\tau)\right\} = \underline{R}(t)\,\delta(t-\tau) \tag{2.8}$$

The basic problem is to estimate the state $\underline{x}(t)$ (or $\underline{z}(t)$) by observing

the signal

$$\underline{w}(\tau) = \underline{y}(\tau) + \underline{v}(\tau) \tag{2.9}$$

over a period of time

$$t_o \leq \tau \leq t \tag{2.10}$$

3. SOME MATHEMATICAL RELATIONS

In this section we shall state some self-evident mathematical relations that relate the two system representations S_1 and S_2.

First we start with the algebraic relation (2.3), i.e.

$$\underline{z} = \underline{g}(\underline{x}) \tag{3.1}$$

We differentiate with respect to time and use the chain rule to obtain

$$\dot{\underline{z}} = \left(\frac{\partial \underline{g}}{\partial \underline{x}}\right) \dot{\underline{x}} = \left(\frac{\partial \underline{g}}{\partial \underline{x}}\right) \underline{f}_1(\underline{x}) \tag{3.2}$$

where $(\partial \underline{g} / \partial \underline{x})$ is the Jacobian matrix of $\underline{g}(.)$. Now since

$$\dot{\underline{z}} = \underline{f}_2(\underline{z}) \tag{3.3}$$

then, in view of Eq. (3.1), we have

$$\dot{\underline{z}} = \underline{f}_2(\underline{g}(\underline{x})) \tag{3.4}$$

From Eqs. (3.2) and (3.4) we deduce the relation

$$\left(\frac{\partial \underline{g}}{\partial \underline{x}}\right) \underline{f}_1(\underline{x}) = \underline{f}_2(\underline{g}(\underline{x})) \quad \text{for all } \underline{x} \tag{3.5}$$

Similarly, from Eqs (2.2) and (2.6) we obtain, in view of Eq. (3.1), the relation

$$\underline{h}_1(\underline{x}) = \underline{h}_2(\underline{g}(\underline{x})) \quad \text{for all } \underline{x} \tag{3.6}$$

Let us now define the matrix $\underline{G}(\underline{x})$ by

$$\underline{G}(\underline{x}) \triangleq \left(\frac{\partial \underline{g}}{\partial \underline{x}}\right) \tag{3.7}$$

and the matrix $\underline{\Gamma}(z)$ (the Jacobian matrix of $g^{-1}(\underline{z})$) by

$$\underline{\Gamma}(\underline{z}) \triangleq \left(\frac{\partial \underline{g}^{-1}}{\partial \underline{z}}\right) \tag{3.8}$$

Since

$$\dot{\underline{z}} = \underline{G}(\underline{x})\,\dot{\underline{x}} \tag{3.9}$$

and

$$\dot{\underline{x}} = \underline{\Gamma}(z)\,\dot{\underline{z}} \tag{3.10}$$

it follows that

$$\underline{G}(\underline{x})\,\underline{\Gamma}(\underline{z}) = \underline{I} \quad \text{for all } \underline{x} \text{ and } \underline{z}\,. \tag{3.11}$$

where $\underline{I}$ is the identity matrix.

4. DYNAMIC LINEARIZATION

Let t_o be some initial time. Let

$$\underline{x}^*(t_o) \tag{4.1}$$

denote some nominal initial state for the system S_1. Let

$$\underline{z}^*(t_o) = \underline{g}(\underline{x}^*(t_o)) \tag{4.2}$$

denote the nominal initial state for the system S_2. Let

$$\underline{x}^*(t) \quad ; \quad t \geq t_o \tag{4.3}$$

denote the nominal trajectory (solution) of the differential equation (2.1) and let $\underline{z}^*(t)$, given by,

$$\underline{z}^*(t) = \underline{g}(\underline{x}^*(t)) \tag{4.4}$$

denote the nominal trajectory (solution) of the system S_2.

We shall denote by $\underline{\xi}(t)$ small perturbations about the trajectory $\underline{x}^*(t)$ so that the actual state of S_1, $\underline{x}(t)$, is given by

$$\underline{x}(t) = \underline{x}^*(t) + \underline{\xi}(t) \tag{4.5}$$

We shall denote by $\underline{J}(t)$ small perturbations about the trajectory $\underline{z}^*(t)$ so that the actual state of S_2, $\underline{z}(t)$, is given by

$$\underline{z}(t) = \underline{z}^*(t) + \underline{J}(t) \tag{4.6}$$

Then for the system S_1 we have

$$\dot{\underline{x}}^*(t) + \dot{\underline{\xi}}(t) = \underline{f}_1(\underline{x}^*(t)) + \left(\frac{\partial \underline{f}_1}{\partial \underline{x}}\right)_{\underline{x}^*(t)} \underline{\xi}(t) + \underline{o}(\underline{\xi}) \tag{4.7}$$

where $\underline{o}(\underline{\xi})$ are the (vector-valued) higher order terms. We define the matrix $\underline{F}_1(\underline{x})$

$$\underline{F}_1(\underline{x}) \triangleq \left(\frac{\partial \underline{f}_1}{\partial \underline{x}}\right) \tag{4.8}$$

and the matrix

$$\underline{F}_1^*(t) \triangleq \underline{F}_1(\underline{x}^*(t)) \triangleq \left(\frac{\partial \underline{f}_1}{\partial \underline{x}}\right)_{\underline{x}^*(t)} \tag{4.9}$$

By neglecting the higher order terms in Eq. (4.7) we have

$$\dot{\underline{\xi}} \cong \underline{F}_1^*(t)\,\underline{\xi} \tag{4.10}$$

In a similar manner we have for the system S_2

$$\dot{\underline{z}}^*(t) + \dot{\underline{J}}(t) = \underline{f}_2(\underline{z}^*(t)) + \left(\frac{\partial \underline{f}_2}{\partial \underline{z}}\right)_{\underline{z}^*(t)} \underline{J}(t) + \underline{o}(J) \tag{4.11}$$

Define the matrices

$$\underline{F}_2(\underline{z}) \triangleq \left(\frac{\partial \underline{f}_2}{\partial \underline{z}}\right) \tag{4.12}$$

and

$$\underline{F}_2^*(t) \triangleq \underline{F}_2(\underline{z}^*(t)) \triangleq \left(\frac{\partial \underline{f}_2}{\partial \underline{z}}\right)_{\underline{z}^*(t)} \tag{4.13}$$

By neglecting the higher order terms $\underline{o}(J)$ in Eq. (4.11) we have

$$\dot{\underline{J}} \cong \underline{F}_2^*(t)\,\underline{J} \tag{4.14}$$

We let $\underline{y}^*(t)$ denote the nominal output of the system. Let $\underline{\eta}(t)$ denote the perturbations about the nominal output so that the actual output $\underline{y}(t)$, for both systems, is

$$\underline{y}(t) = \underline{y}^*(t) + \underline{\eta}(t) \tag{4.15}$$

From Eqs. (2.1) and (4.15) we have

$$\underline{y}^*(t) + \underline{\eta}(t) = \underline{h}_1(\underline{x}^*(t)) + \left(\frac{\partial \underline{h}_1}{\partial \underline{x}}\right)_{\underline{x}^*(t)} \underline{\xi}(t) + \underline{o}_1(\underline{\xi}) \tag{4.16}$$

while from Eqs. (2.6) and (4.15) we have

$$\underline{y}^*(t) + \underline{\eta}(t) = \underline{h}_2(\underline{z}^*(t)) + \left(\frac{\partial \underline{h}_2}{\partial \underline{z}}\right)_{\underline{z}^*(t)} \underline{J}(t) + \underline{o}_1(J) \tag{4.17}$$

Define the matrices:

$$\underline{H}_1(\underline{x}) \triangleq \left(\frac{\partial \underline{h}_1}{\partial \underline{x}}\right) \tag{4.18}$$

$$\underline{H}_1^*(t) \triangleq \underline{H}_1(\underline{x}^*(t)) \triangleq \left(\frac{\partial \underline{h}_1}{\partial \underline{x}}\right)_{\underline{x}^*(t)} \tag{4.19}$$

$$\underline{H}_2(\underline{z}) \triangleq \left(\frac{\partial \underline{h}_2}{\partial \underline{z}}\right) \tag{4.20}$$

$$\underline{H}_2^*(t) \triangleq \underline{H}_2(\underline{z}^*(t)) \triangleq \left(\frac{\partial \underline{h}_2}{\partial \underline{z}}\right)_{\underline{z}^*(t)} \tag{4.21}$$

Clearly

$$\underline{\eta}(t) = \underline{H}_1^*(t)\,\underline{\xi}(t) + \underline{o}_1(\underline{\xi}) = \underline{H}_2^*(t)\,\underline{J}(t) + \underline{o}_1(J) \tag{4.22}$$

If we neglect the higher order terms $\underline{o}_1(\underline{\xi})$ and $\underline{o}_1(J)$ we have

$$\underline{\eta}(t) \cong \underline{H}_1^*(t)\,\underline{\xi}(t) \tag{4.23}$$

$$\underline{\eta}(t) \cong \underline{H}_2^*(t)\,\underline{J}(t) \tag{4.24}$$

Recall that (see Eq. (2.29)) the observed signal is

$$\underline{\omega}(\tau) = \underline{y}(\tau) + \underline{v}(\tau) \tag{4.25}$$

so that

$$\underline{\omega}(\tau) = \underline{y}^*(\tau) + \underline{\eta}(\tau) + \underline{v}(\tau) \tag{4.26}$$

Since $\underline{y}^*(\tau)$ is a deterministic and computable signal, we can think of the observed signal as being

$$\underline{\omega}(\tau) - \underline{y}^*(\tau) = \underline{\eta}(\tau) + \underline{v}(\tau) \tag{4.27}$$

The above linearization procedure yields the following results. The linearization of the nonlinear system with $\underline{x}$ as the state vector leads to a linear and time-varying system, L_1, with state vector $\underline{\xi}$ and output vector $\underline{\eta}$ related by

$$
L_1 \begin{cases} \dot{\underline{\xi}} = \underline{F}_1^*(t)\,\underline{\xi} & (4.28) \\ \underline{\eta} = \underline{H}_1^*(t)\,\underline{\xi} & (4.29) \end{cases}
$$

On the other hand, the linearization of the nonlinear system with $\underline{z}$ as the state vector leads to a linear and time-varying system, L_2, with state vector $\underline{J}$ and output vector $\underline{\eta}$ related by

$$
L_2 \begin{cases} \dot{\underline{J}} = \underline{F}_2^*(t)\,\underline{J} & (4.30) \\ \underline{\eta} = \underline{H}_2^*(t)\,\underline{J} & (4.31) \end{cases}
$$

We shall now proceed to derive the relations which relate the matrices $\underline{F}_1^*(t)$, $\underline{H}_1^*(t)$, $\underline{G}^*(t)$, $\underline{F}_2^*(t)$, and $\underline{H}_2^*(t)$ in Eqs. (4.28) through (4.31). To do this, we start from Eq. (4.4), namely,

$$
\underline{z}^*(t) = \underline{g}(\underline{x}^*(t)) \tag{4.32}
$$

Therefore,

$$
\underline{z}^*(t) + \underline{J}(t) = \underline{g}(\underline{x}^*(t) + \underline{\xi}(t) \tag{4.33}
$$

$$
= \underline{g}(\underline{x}^*(t)) + \left(\frac{\partial \underline{g}}{\partial \underline{x}}\right)_{\underline{x}^*(t)} \xi(t) + \underline{o}_2(\xi) \tag{4.34}
$$

Let (see Eq. (3.7))

$$
\underline{G}^*(t) \triangleq \underline{G}(\underline{x}^*(t)) \triangleq \left(\frac{\partial \underline{g}}{\partial \underline{x}}\right)_{\underline{x}^*(t)} \tag{4.35}
$$

Then, from Eqs (4.35), (4.34), and (4.32) we deduce the relation

$$
\underline{J}(t) = \underline{G}^*(t)\,\underline{\xi}(t) + \underline{o}_2(\underline{\xi}) \tag{4.36}
$$

Neglecting the higher order terms $\underline{o}_2(\underline{\xi})$ we have for small $\|\underline{\xi}(t)\|$

$$\underline{J}(t) = \underline{G}^*(t)\,\underline{\xi}(t) \tag{4.37}$$

Differentiating with respect to time we have

$$\dot{\underline{J}}(t) = \dot{\underline{G}}^*(t)\,\underline{\xi}(t) + \underline{G}^*(t)\,\dot{\underline{\xi}}(t) \tag{4.38}$$

Substituting Eqs. (4.30) and (4.28) into Eq. (4.38) we obtain

$$\underline{F}_2^*(t)\,\underline{J}(t) = \left[\dot{\underline{G}}^*(t) + \underline{G}^*(t)\,\underline{F}_1^*(t)\right]\,\underline{\xi}(t) \tag{4.39}$$

Substituting Eq. (4.37) into Eq. (4.39) and noting that the ensuing relation must hold for all $\underline{\xi}(t)$ we deduce that

$$\underline{F}_2^*(t)\,\underline{G}^*(t) = \dot{\underline{G}}^*(t) + \underline{G}^*(t)\,\underline{F}_1^*(t) \tag{4.40}$$

or

$$\underline{F}_2^*(t) = \dot{\underline{G}}^*(t)\,\underline{G}^{*-1}(t) + \underline{G}^*(t)\,\underline{F}_1^*(t)\,\underline{G}^{*-1}(t) \tag{4.41}$$

Clearly, Eq. (4.41) links the matrices $\underline{F}_2^*(t)$ and $\underline{F}^*(t)$ appearing in Eqs. (4.28) and (4.30).

Next, we proceed to relate the matrices $\underline{H}_1^*(t)$ and $\underline{H}_2^*(t)$. From Eqs. (4.29), (4.31), and (4.37) we conclude that

$$\underline{H}_2^*(t) = \underline{H}_1^*(t)\,\underline{G}^{*-1}(t) \tag{4.42}$$

Let

$$\underline{\Gamma}^*(t) \triangleq \underline{G}^{*-1}(t) \tag{4.43}$$

Then Eqs. (4.41) and (4.42) reduce to

$$\underline{F}_2^*(t) = \dot{\underline{G}}^*(t)\,\underline{\Gamma}^*(t) + \underline{G}^*(t)\,\underline{F}_1^*(t)\,\underline{\Gamma}^*(t) \tag{4.44}$$

$$\underline{H}_2^*(t) = \underline{H}_1^*(t)\,\underline{\Gamma}^*(t) \tag{4.45}$$

We remark that the above equations hold if and only if the various higher order terms are neglected.

5. THE KALMAN FILTERS AND THE COVARIANCE EQUATIONS

Let us consider the linear system L_1 defined by Eqs. (4.28) and (4.29). The observed signal is (see Eq. (4.27))

$$\underline{\tilde{\omega}}(t) = \underline{\eta}(t) + \underline{v}(t) \tag{5.1}$$

We can construct a Kalman-Bucy filter (see Ref. [1]) which will yield an estimate $\hat{\underline{\xi}}(t)$ of $\underline{\xi}(t)$. The filter is specified by the equation

$$\frac{d}{dt}\hat{\underline{\xi}}(t) = \left[\underline{F}_1^{*}(t) - \underline{\Sigma}_\xi(t)\,\underline{H}_1^{*\prime}(t)\,\underline{R}^{-1}(t)\,\underline{H}_1^{*}(t)\right]\hat{\underline{\xi}}(t)$$

$$+ \underline{\Sigma}_\xi(t)\,\underline{H}_1^{*\prime}(t)\,\underline{R}^{-1}(t)\,\underline{\tilde{\omega}}(t) \quad ; \quad \hat{\underline{\xi}}(t_o) = \underline{0} \tag{5.2}$$

where $\underline{\Sigma}_\xi(t)$ is the covariance matrix of the error vector $\underline{\xi}(t) - \hat{\underline{\xi}}(t)$ and it satisfies the Riccati matrix differential equation

$$\frac{d}{dt}\underline{\Sigma}_\xi = \underline{F}_1^{*}(t)\,\underline{\Sigma}_\xi + \underline{\Sigma}_\xi\,\underline{F}_1^{*\prime}(t) - \underline{\Sigma}_\xi\underline{H}_1^{*\prime}(t)\,\underline{R}^{-1}(t)\,\underline{H}_1^{*}(t)\,\underline{\Sigma}_\xi \tag{5.3}$$

Similarly consider the linear system L_2 defined by Eqs. (4.30) and (4.31) with the observed signal given by Eq. (5.1). Again we can construct a Kalman-Bucy filter which generates an estimate $\hat{\underline{J}}(t)$ of $\underline{J}(t)$. The filter is specified by

$$\frac{d}{dt}\hat{\underline{J}}(t) = \left[\underline{F}_2^{*}(t) - \underline{\Sigma}_J(t)\,\underline{H}_2^{*\prime}(t)\,\underline{R}^{-1}(t)\,\underline{H}_2^{*}(t)\right]\hat{\underline{J}}(t)$$

$$+ \underline{\Sigma}_J(t)\,\underline{H}_2^{*\prime}(t)\,\underline{R}^{-1}(t)\,\underline{\tilde{\omega}}(t) \quad ; \quad \hat{\underline{J}}(t_o) = \underline{0} \tag{5.4}$$

where $\underline{\Sigma}_J(t)$ is the covariance matrix of the error vector $\underline{J}(t) - \hat{\underline{J}}(t)$ and it satisfies the Riccati equation

$$\frac{d}{dt}\underline{\Sigma}_J = \underline{F}_2^{*}(t)\,\underline{\Sigma}_J + \underline{\Sigma}_J\,\underline{F}_2^{*\prime}(t) - \underline{\Sigma}_J\,\underline{H}_2^{*\prime}(t)\,\underline{R}^{-1}(t)\,\underline{H}_2^{*}(t)\,\underline{\Sigma}_J \tag{5.5}$$

We shall now show that the covariance matrices $\underline{\Sigma}_\xi(t)$ and $\underline{\Sigma}_J(t)$ are related by the relation

$$\underline{\Sigma}_{\tilde{x}}(t) = \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\Gamma^{*\prime}(t) \tag{5.6}$$

To do this we proceed as follows. Let

$$\underline{X}(t) \triangleq \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{\Gamma}^{*\prime}(t) \tag{5.7}$$

Then,

$$\begin{aligned}\frac{d}{dt}\underline{X}(t) &= \dot{\underline{\Gamma}}^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{\Gamma}^{*\prime}(t) \\ &+ \underline{\Gamma}^{*}(t)\,\dot{\underline{\Sigma}}_J(t)\,\underline{\Gamma}^{*\prime}(t) \\ &+ \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\dot{\underline{\Gamma}}^{*\prime}(t)\end{aligned} \tag{5.8}$$

Substituting Eq. (5.5) into Eq. (5.8) we obtain

$$\begin{aligned}\dot{\underline{X}}(t) &= \dot{\underline{\Gamma}}^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{\Gamma}^{*\prime}(t) + \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\dot{\underline{\Gamma}}^{*\prime}(t) \\ &+ \underline{\Gamma}^{*}(t)\,\underline{F}_2^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{\Gamma}^{*\prime}(t) + \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{F}_2^{*\prime}(t)\,\underline{\Gamma}^{*\prime}(t) \\ &- \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{H}_2^{*\prime}(t)\,\underline{R}^{-1}(t)\,\underline{H}_2^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{\Gamma}^{*\prime}(t)\end{aligned} \tag{5.9}$$

Next we substitute Eqs. (4.44) and (4.45) into Eq. (5.9); we use the relation that (since from Eq. (4.43), $\underline{\Gamma}^{*}(t)\,\underline{G}^{*}(t) = \underline{I}$)

$$\underline{\Gamma}^{*}(t) = -\underline{\Gamma}^{*}(t)\,\dot{\underline{G}}^{*}(t)\,\underline{\Gamma}^{*}(t) \tag{5.10}$$

to obtain

$$\begin{aligned}\dot{\underline{X}}(t) &= -\underline{\Gamma}^{*}(t)\,\dot{\underline{G}}^{*}(t)\,\underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{\Gamma}^{*\prime}(t) \\ &- \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{\Gamma}^{*\prime}(t)\,\dot{\underline{G}}^{*\prime}(t)\,\underline{\Gamma}^{*\prime}(t) \\ &+ \underline{\Gamma}^{*}(t)\,\dot{\underline{G}}^{*}(t)\,\underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\Gamma^{*\prime}(t) \\ &+ \underline{\Gamma}^{*}(t)\,\underline{G}^{*}(t)\,\underline{F}^{*}(t)\,\underline{\Gamma}^{*}(t)\,\underline{\Sigma}_J(t)\,\underline{\Gamma}^{*\prime}(t)\end{aligned}$$

$$+ \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_{J}(t)\,\underline{\Gamma}^{*'}(t)\,\dot{\underline{G}}^{*'}(t)\,\underline{\Gamma}^{*'}(t)$$

$$+ \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_{J}(t)\,\underline{\Gamma}^{*'}(t)\,\underline{F}^{*'}(t)\,\underline{G}^{*'}(t)\,\underline{\Gamma}^{*'}(t)$$

$$- \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_{J}(t)\,\underline{\Gamma}^{*'}(t)\,\underline{H}_{1}^{*'}(t)\,\underline{R}^{-1}(t)\,\underline{H}_{1}^{*}(t)\,\underline{\Gamma}^{*}(t)\,\underline{\Sigma}_{J}(t)\,\underline{\Gamma}^{*'}(t) \qquad (5.\,11)$$

Substituting Eq. (5. 7) into Eq. (5. 11) we obtain

$$\dot{\underline{X}}(t) = + \underline{X}(t)\,\underline{F}_{1}^{*'}(t)\,\underline{G}^{*'}(t)\,\underline{\Gamma}^{*'}(t)$$

$$- \underline{X}(t)\,\underline{H}_{1}^{*'}(t)\,\underline{R}^{-1}(t)\,\underline{H}_{1}^{*}(t)\,\underline{X}(t) \qquad (5.\,12)$$

Recalling that $\underline{\Gamma}^{*}(t)\,\underline{G}^{*}(t) = \underline{I}$ we see that Eq. (5. 12) reduces further to

$$\dot{\underline{X}}(t) = \underline{F}_{1}^{*}(t)\,\underline{X}(t) + \underline{X}(t)\,\underline{F}_{1}^{*'}(t) - \underline{X}(t)\,\underline{H}_{1}^{*'}(t)\,\underline{R}^{-1}(t)\,\underline{H}_{1}^{*}(t)\,\underline{X}(t) \qquad (5.\,13)$$

Let us now compare Eqs. (5. 13) and (5. 3). We conclude that $\underline{X}(t)$ and $\underline{\Sigma}_{\xi}(t)$ satisfy the <u>same</u> matrix differential equations. Furthermore, from Eqs (5. 7) and (4. 37), $\underline{X}(t_o) = \underline{\Gamma}^{*}(t_o)\,\underline{\Sigma}_{J}(t_o)\,\underline{\Gamma}^{*'}(t_o)$

$$\underline{X}(t_o) = \underline{\Gamma}^{*}(t_o)\,\underline{\Sigma}_{J}(t_o)\,\underline{\Gamma}^{*'}(t_o)$$

$$= \underline{\Gamma}^{*}(t_o)\; E\left\{\underline{J}(t_o)\,\underline{J}'(t_o)\right\}\;\underline{\Gamma}^{*'}(t_o)$$

$$= \underline{\Gamma}^{*}(t_o)\,\underline{G}^{*}(t_o)\; E\left\{\underline{\xi}(t_o)\,\underline{\xi}'(t_o)\right\}\;\underline{G}^{*'}(t_o)\,\underline{\Gamma}^{*'}(t_o)$$

$$= E\left\{\underline{\xi}(t_o)\,\underline{\xi}'(t_o)\right\} \triangleq \underline{\Sigma}_{\xi}(t_o) \qquad (5.\,14)$$

so that $\underline{X}$ and $\underline{\Sigma}_{\xi}$ have the same initial conditions. Therefore, by the uniqueness theorem of solutions to differential equations we conclude that

$$\underline{X}(t) = \underline{\Sigma}_{\xi}(t) \qquad \text{for all } t \qquad (5.\,15)$$

and, hence, from Eq. (5. 7), that

$$\underline{\Sigma}_{\xi}(t) = \underline{\Gamma}^{*}(t)\,\underline{\Sigma}_{J}(t)\,\underline{\Gamma}^{*'}(t) \qquad (5.\,16)$$

Up to now we have shown that the two error covariances matrices $\underline{\Sigma}_{\xi}(t)$ and $\underline{\Sigma}_{J}(t)$ are related by the linear equation (5.16); this can also be written as

$$\underline{\Sigma}_{J}(t) = \underline{G}^{*}(t)\underline{\Sigma}_{\xi}(t)\underline{G}^{*\prime}(t) \tag{5.16}$$

The next step is to compare the estimates $\hat{\underline{\xi}}(t)$ and $\hat{\underline{J}}(t)$ of the two Kalman filters (5.2) and (5.4). We shall show that

$$\hat{\underline{J}}(t) = \underline{G}^{*}(t)\hat{\underline{\xi}}(t) \tag{5.17}$$

To do this, define the vector

$$\underline{q}(t) \triangleq \underline{\Gamma}^{*}(t)\hat{\underline{J}}(t) \tag{5.18}$$

Note that

$$\underline{q}(t_o) = \underline{\Gamma}^{*}(t_o)\hat{\underline{J}}(t_o) = \underline{0} = \hat{\underline{\xi}}(t_o) \tag{5.19}$$

We shall show that $\underline{q}(t)$ and $\hat{\underline{\xi}}(t)$ satisfy the same differential equation; also, by (5.19) they have identical initial conditions; this would imply that $\underline{q}(t) = \hat{\underline{\xi}}(t)$ and, therefore, Eq. (5.17) would follow.

From Eq. (5.18) we have

$$\frac{d}{dt}\underline{q}(t) = \dot{\underline{\Gamma}}^{*}(t)\hat{\underline{J}}(t) + \underline{\Gamma}^{*}(t)\frac{d}{dt}\hat{\underline{J}}(t) \tag{5.20}$$

$$= -\underline{\Gamma}^{*}(t)\dot{\underline{G}}^{*}(t)\underline{\Gamma}^{*}(t)\hat{\underline{J}}(t) + \underline{\Gamma}^{*}(t)\frac{d}{dt}\hat{\underline{J}}(t) \tag{5.21}$$

$$= -\underline{\Gamma}^{*}(t)\dot{\underline{G}}^{*}(t)\underline{q}(t) \tag{5.22}$$

$$+\left[\underline{\Gamma}^{*}(t)\underline{F}_2^{*}(t) - \underline{\Gamma}^{*}(t)\underline{\Sigma}_J(t)\underline{H}_2^{*\prime}(t)\underline{R}^{-1}(t)\underline{H}_2^{*}(t)\right]\hat{\underline{J}}(t)$$

$$+\underline{\Gamma}^{*}(t)\underline{\Sigma}_J(t)\underline{H}_2^{*\prime}(t)\underline{R}^{-1}(t)\tilde{\underline{\omega}}(t) \tag{5.23}$$

Substituting Eqs. (4.44), (4.45), and (5.16) into Eq. (5.23) we obtain after some algebraic manipulations

$$\frac{d}{dt}\underline{q}(t) = \left[\underline{F}_1^*(t) - \underline{\Sigma}_\xi(t)\,\underline{H}_1^{*'}(t)\,\underline{R}^{-1}(t)\,\underline{H}_1^*(t)\right]\underline{q}(t)$$

$$+ \underline{\Sigma}_\xi(t)\,\underline{H}_1^{*'}(t)\,\underline{R}^{-1}(t)\,\underline{\tilde{\omega}}(t) \qquad (5.24)$$

Comparing Eqs. (5.24) and (5.2) we deduce that $\hat{\underline{\xi}}(t)$ and $\underline{q}(t)$ satisfy the same differential equation. Thus, on the basis of our previous argument we conclude that

$$\hat{\underline{J}}(t) = \underline{G}^*(t)\,\hat{\underline{\xi}}(t) \qquad (5.25)$$

The implications of this result will be discussed in the following section.

6. Conclusions

Let us recapitulate the development and the results up to now.

We started with two different state variable representations of the same system (see Eqs. (2.1), (2.2) and (2.5), (2.6)). We linearized the system about the same nominal trajectory. In this manner we derived two linear systems, describing the equations of the perturbation vectors, L_1 and L_2 (see Eqs. (4.28) through (4.31)). These systems were described by

$$\left.\begin{aligned} \dot{\underline{\xi}}(t) &= \underline{F}^*(t)\,\underline{\xi}(t) \\ \dot{\underline{J}}(t) &= \underline{A}^*(t)\,\underline{J}(t) \end{aligned}\right\} \tag{6.1}$$

We showed that (see Eq. (4.37))

$$\underline{J}(t) = \underline{G}^*(t)\,\underline{\xi}(t) \tag{6.2}$$

to within first order terms where $\underline{G}^*(t)$ is a computable matrix from the problem data (see Eq. (4.35)).

Next we showed how to generate the estimate $\hat{\underline{\xi}}(t)$ of $\underline{\xi}(t)$ and the estimate $\hat{\underline{J}}(t)$ of $\underline{J}(t)$ using the Kakman-Bucy Filter. After many manipulations we showed that the estimates $\hat{\underline{\xi}}(t)$ and $\hat{\underline{J}}(t)$ were related by (see Eq. (5.25))

$$\hat{\underline{J}}(t) = \underline{G}^*(t)\,\hat{\underline{\xi}}(t) \tag{6.3}$$

In other words, the estimates are related by the same linear transformation, $\underline{G}^*(t)$, as the perturbation vectors. Since $\underline{G}^*(t)$ is a deterministic matrix we conclude that we can construct a Kalman Filter based upon the linearized state equations in any one coordinate system to obtain the estimate, $\hat{\underline{x}}(t)$, of the state $\underline{x}(t)$ in that coordinate system by

$$\hat{\underline{x}}(t) = \underline{x}^*(t) + \hat{\underline{\xi}}(t) \tag{6.4}$$

where $\underline{x}^*(t)$ is the deterministic nominal trajectory. If we chose a different coordinate system in which the state is $\underline{z}(t)$ and given that $\underline{z}$ and $\underline{x}$ are related by

$$\underline{z} = \underline{g}(x) \quad \text{or} \quad \underline{x} = \underline{\gamma}(\underline{z}) \tag{6.5}$$

Then if we let

$$\underline{G}^*(t) \triangleq \left(\frac{\partial \underline{g}}{\partial \underline{x}} \right)_{\underline{x}^*(t)} \tag{6.6}$$

then we can construct the estimate $\hat{\underline{z}}(t)$ by

$$\hat{\underline{z}}(t) = \underline{g}(\underline{x}^*(t)) + \underline{G}^*(t)\,\hat{\underline{\xi}}(t) \tag{6.7}$$

using the output of the previous Kalman filter. If higher order terms in the linearization scheme are negligible, the estimate $\hat{\underline{z}}(t)$ is optimal.

The above discussion does not mean that there need not exist a coordinate system in which the linearization is more accurate than in another coordinate system. However, it is not known how to find such a coordinate system.

ACKNOWLEDGMENT

The author expresses his sincere thanks to Dr. Richard P. Wishner and Mr. Michael Gruber for their helpful comments and discussions.

7. REFERENCES

1. R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," J. of Basic Engineering, Trans. ASME, 83D, 1961, pp 95-108.

DISTRIBUTION

Director's Office

Division I

R. E. Rader

Division II

F. C. Frick

Group 21

K. E. Ralston

Group 22

W. Z. Lemnios
D. H. Pruslin

Division III

S. H. Dodd, Jr.

Group 35

K. Kresa

Division IV

J. Freedman
G. F. Watson
H. G. Weiss

Group 41

D. L. Clark
H. M. Jones
J. B. Resnick

Group 42

M. E. Austin
A. Bertolini
D. R. Bromaghim
G. F. Dalrymple
D. F. DeLong
A. A. Galvin
H. M. Heggestad
E. M. Hofstetter
O. A. Z. Leneman
B. Loesch
J. Margolin
G. R. McCully
J. P. Perry
C. M. Steinmetz
J. A. Tabaczynski
J. O. Taylor
R. P. Wishner (10)

Group 43

R. C. Butman

Group 44

C. Blake

Group 45

W. W. Camp
V. L. Lynn
S. J. Miller

Group 46

C. W. Jones
C. E. Muehe

Division VI

W. E. Morrow, Jr.

Group 62

K. L. Jordan
I. Stiglitz

Group 64

P. E. Green
R. T. Lacoss

Group 66

T. J. Goblick
B. Reiffen

Division VII

J. F. Hutzenlaub

Division VIII

A. L. McWhorter

DISTRIBUTION

Massachusetts Institute of Technology

Prof. M. Athans (20)	Room	24-306
Prof. R. W. Brockett	"	24-302
Mr. J. Burchfiel	"	24-316
Mr. R. Canales	"	24-316
Mr. A. Debs	"	24-316
Prof. M. Dertovzos	"	32-100
Prof. L. A. Gould	"	32-102
Mr. D. L. Gray	"	24-316
Mr. S. Greenberg	"	24-316
Mr. D. L. Kleinman	"	24-316
Mr. W. Levine	"	24-316
Mr. A. Levis	"	24-316
Prof. G. C. Newton	"	32-106
Prof. F. C. Schweppe	"	24-308

Massachusetts Institute of Technology
Department of Electrical Engineering

Prof. W. Davenport
Prof. A. Drake
Prof. R. Gallager
Prof. R. Kennedy
Prof. L. D. Smullin
Prof. H. L. Snyder
Prof. H. VanTrees

DOCUMENT CONTROL DATA – R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Lincoln Laboratory, M.I.T.	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP None	
3. REPORT TITLE On the Equivalence of Linearized Kalman Filters		
4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)* Technical Note		
5. AUTHOR(S) *(Last name, first name, initial)* Athans, Michael		
6. REPORT DATE 11 January 1967	7a. TOTAL NO. OF PAGES 28	7b. NO. OF REFS 1
8a. CONTRACT OR GRANT NO. AF 19(628)-5167 b. PROJECT NO. ARPA 498 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Note 1967-6 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* ESD-TR-67-20	
10. AVAILABILITY/LIMITATION NOTICES Distribution of this document is unlimited.		
11. SUPPLEMENTARY NOTES None	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency, Department of Defense	

13. ABSTRACT

The purpose of this note is to examine the relationships of the Kalman filters which can be used to estimate the linearized state of two different state-variable representations of the same nonlinear system. It is shown that, to within first order effects, the choice of the coordinate system for the construction of the Kalman filter is immaterial.

14. KEY WORDS

Kalman filters
linearized filters
nonlinear systems
trajectories